Contrôle de connaissances SD211 (2016/2017)

Documents autorisés : polycopié de l'école et notes personnelles de l'élève.
Presque toutes les questions peuvent être traitées indépendamment.

Notation : ι_A désigne l'indicatrice convexe d'un ensemble A.

On considère un système comportant N générateurs d'électricité indépendants où $N \geq 1$ est un entier. Chaque générateur n dispose d'une fonction $f_n : \mathbb{R} \to [0, +\infty]$, convexe propre et fermée, telle que $f_n(x)$ représente le coût en euros nécessaire à ce générateur pour produire x Watt-heure, pour tout $x \geq 0$ (on posera $f_n(x) = +\infty$ pour $x < 0$). Le système doit pouvoir satisfaire une demande globale égale à D, c'est à dire que, si x_n représente la production du $n$ème générateur, la contrainte $\sum_{n=1}^N x_n \geq D$ doit être vérifiée. On cherche à minimiser le coût global, c'est à dire qu'on se pose le problème

$$\min_{\boldsymbol{x} \in \mathbb{R}^N : \sum_{n=1}^N x_n \geq D} \sum_{n=1}^N f_n(x_n)\,. \tag{1}$$

1. Ecrire le lagrangien associé au problème (1).
2. Montrer que la fonction duale Φ satisfait pour tout $\nu \geq 0$,

$$\Phi(\nu) = -\sum_{n=1}^N f_n^*(\nu) + \nu\, D \tag{2}$$

   où f_n^* est la transformée de Legendre de f_n.
3. Montrer que le problème dual est équivalent à

$$\min_{\nu \geq 0} \sum_{n=1}^N f_n^*(\nu) - \nu D\,. \tag{3}$$

On cherche à résoudre le problème dual de manière distribuée : chaque générateur n est un *agent* disposant d'un calculateur. Soit $\mathcal{C}$ l'espace vectoriel défini par

$$\mathcal{C} = \left\{ \boldsymbol{\nu} \in \mathbb{R}^N : \nu_1 = \cdots = \nu_N \right\}\,.$$

On pose pour tout $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_N)$,

$$F(\boldsymbol{\nu}) = \sum_{n=1}^N f_n^*(\nu_n) \quad \text{et} \quad G(\boldsymbol{\nu}) = \iota_{\mathcal{C}}(\boldsymbol{\nu}) + \iota_{\mathbb{R}_+}(\nu_1) - \nu_1 D\,.$$

On cherche dorénavant à résoudre le problème d'optimisation

$$\min_{\boldsymbol{\nu} \in \mathbb{R}^N} F(\boldsymbol{\nu}) + G(\boldsymbol{\nu}) \tag{4}$$

4. Quel lien y a-t-il entre les solutions du problème (4) et les solutions du problème dual (3) ?
5. Pour résoudre le problème (4), on met en oeuvre la méthode suivante, où $\rho > 0$ est un paramètre et où $k \in \mathbb{N}$ est l'indice de l'itération. Compléter les pointillés dans l'équation (7)

$$\boldsymbol{\nu}^{k+1} = \arg\min_{\boldsymbol{\nu}} F(\boldsymbol{\nu}) + \langle \boldsymbol{\lambda}^k, \boldsymbol{\nu} \rangle + \frac{1}{2\rho} \|\boldsymbol{\nu} - \boldsymbol{\omega}^k\|^2 \tag{5}$$

$$\boldsymbol{\omega}^{k+1} = \arg\min_{\boldsymbol{\omega}} G(\boldsymbol{\omega}) - \langle \boldsymbol{\lambda}^k, \boldsymbol{\omega} \rangle + \frac{1}{2\rho} \|\boldsymbol{\nu}^{k+1} - \boldsymbol{\omega}\|^2 \tag{6}$$

$$\boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k + \cdots \tag{7}$$

   Quel nom porte cet algorithme ?
6. On cherche à résoudre explicitement l'arg min dans l'équation (6). On définit $\bar{\lambda}^k = \frac{1}{N} \sum_n \boldsymbol{\lambda}_n^k$. Montrer que la solution s'écrit $\boldsymbol{\omega}^{k+1} = (u_{k+1}, \ldots, u_{k+1})$ où u_{k+1} est l'unique solution du problème de minimisation

$$\min_{u \geq 0} -uN\beta_k + \frac{1}{2\rho} \sum_{n=1}^{N} (u - \nu_n^{k+1})^2 \ , \tag{8}$$

   où β_k est une valeur à exprimer en fonction de D, N et $\bar{\lambda}^k$.
7. Résoudre explicitement le problème (8) et fournir la valeur de u_{k+1} en fonction de β_k, N, ρ et $\bar{\nu}^{k+1}$, où $\bar{\nu}^{k+1} = \frac{1}{N} \sum_n \boldsymbol{\nu}_n^{k+1}$.
8. On s'intéresse maintenant à l'arg min de l'équation (5). En utilisant la question 6, montrer que pour tout $k \geq 1$, la solution est donnée par

$$\boldsymbol{\nu}_n^{k+1} = \arg\min_{\nu \in \mathbb{R}} f_n^*(\nu) + \boldsymbol{\lambda}_n^k \, \nu + \frac{1}{2\rho} (\nu - u^k)^2$$

9. Montrer que $\boldsymbol{\nu}_n^{k+1} = \mathrm{prox}_{\rho f_n^*}(m_k)$ où m_k est à exprimer en fonction de $\boldsymbol{\lambda}_n^k$, ρ et u_k
10. L'architecture logicielle comporte N calculateurs distribués (les agents) communiquant, à chaque itération, à un calculateur central. Seul le calculateur central connait la valeur D. Seul l'agent n connait la valeur de f_n^*. Ecrire le pseudo-code l'algorithme utilisé, en faisant clairement apparaître, à chaque itération :
    - les calculs effectués localement au niveau de chaque agent,
    - les calculs effectués au niveau du calculateur central,
    - les variables transmises par chaque agent au calculateur central, et réciproquement.

Dans la dernière partie de ce sujet, on cherche à mettre en œuvre un autre algorithme distribué que celui proposé à la question 5.

11. En utilisant le cours, donner une condition suffisante pour que, quel que soit $n = 1, \dots, N$, f_n^* soit dérivable et de gradient L_n-lipschitzien, où $L_n > 0$. On suppose désormais que cette condition est satisfaite.
12. On souhaite utiliser l'algorithme du gradient proximal pour résoudre le problème (4). Exprimer en fonction de $L_1, \dots, L_N$ une condition sur le pas de l'algorithme, permettant d'assurer sa convergence.
13. Ecrire les itérations de l'algorithme, avec les mêmes consignes que pour la question 10, c'est à dire en indiquant clairement les calculs effectués localement par les agents et par le calculateur central.